Journal of Bioenergetics and Biomembranes, Vol. 29, No. 3, 1997

Lineage-Specific Evolution of Echinoderm Mitochondrial ATP Synthase Subunit 8

C. De Giorgi,[1] A. Martiradonna,[1] G. Pesole,[2] and C. Saccone[3]

Received July 11, 1996; accepted November 7, 1996

Peculiar evolutionary properties of the subunit 8 of mitochondrial ATP synthase (ATPase8) are revealed by comparative analyses carried out between both closely and distantly related species of echinoderms. The analysis of nucleotide substitution in the three echinoids demonstrated a relaxation of amino acid functional constraints. The deduced protein sequences display a well conserved domain at the N-terminus, while the central part is very variable. At the C-terminus, the broad distribution of positively charged amino acids, which is typical of other organisms, is not conserved in the two different echinoderm classes of the sea urchins and of the sea stars. Instead, a motif of three amino acids, so far not described elsewhere, is conserved in sea urchins and is found to be very similar to the motif present in the sea stars. Our results indicate that the N-terminal region seems to follow the same evolutionary pattern in different organisms, while the maintenance of the C-terminal part in a phylum-specific manner may reflect the co-evolution of mitochondrial and nuclear genes.

KEY WORDS: ATP synthase subunit 8; genes; mammals; mitochondria; sea urchins; sequences.

INTRODUCTION

The animal mitochondrial DNA (mtDNA) consists of a very compact circular DNA molecule containing the genes for a complete set of tRNAs, for two ribosomal RNAs, and for 13 proteins with few exceptions. In spite of this high conservation of gene content, the gene order differs from phylum to phylum, but is relatively constant within the same phylum, where only few transpositions are tolerated. In the phylum echinoderms, only a 4.6-kb fragment transposition differentiates the mt genome of the two classes of Asteroid (sea star) and Echinoid (sea urchin) although their separation can be traced back about 560 million years (Smith, 1984). The complete sequence of the mtDNA of the sea urchin *Arbacia lixula,* order of Stirodonta, has been recently obtained (De Giorgi *et al.,* 1996), thus allowing the comparison of this genome with those of other sea urchin species, *Paracentrotus lividus* and *Strongylocentrotus purpuratus* belonging to the order of Camarodonta, (Jacobs *et al.,* 1988; Cantatore *et al.,* 1989). Molecular evolution studies on the three genomes carried out with the methods developed in our laboratory have indicated that the divergence time between the two orders, Camarodonta and Stirodonta, is between 65 and 155 million years (De Giorgi *et al.,* 1995).

The comparison of the three genomes has revealed the high conservation of gene sequence and organization in the sea urchin lineages. Therefore, unlike other invertebrate mt genomes (such as those of insects and nematodes) where differences are detected also within closely related species (Crozier and Crozier, 1993; Okimoto *et al.,* 1992), mtDNA is highly conserved in echinoids.

In particular, the organization of genes reveals that the tRNA gene cluster, which includes 15 tRNAs and the control region, is located in identical position in the three genomes and is characterized by very relevant compactness. Furthermore the same A-T rich

[1] Dipartimento di Biochimica e Biologia Molecolare, Università di Bari, Via Orabona 4, 70126 Bari, Italy.
[2] Dipartimento di Biologia, D.B.A.F., Università della Basilicata, Via Anzio 10, 85100 Potenza, Italy.
[3] Centro di Studio sui Mitocondri e Metabolismo Energetico, Via Amendola 165/A, 70126 Bari, Italy.

0145-479X/97/0600-0233$12.50/0

palindromic sequences, whose significance is still unknown, are located in corresponding regions in the three genomes (De Giorgi *et al.*, 1995). The comparison between the protein coding genes, with the exception of ATPase8, has revealed that the similarity of the three genomes ranges from 60 to 80% at the nucleotide level, and it increases up to 69–92% at the amino acid level. This observation is taken as an indication of the existence of physiological constraints operating at the amino acid level. The control region itself, which is known to be the least conserved region in the mtDNAs sequenced so far, shows a considerable overall degree of conservation, with similarity ranging from 63.6 to 69.7% between the three sea urchin species.

In this peculiar scenario, the gene for subunit 8 of ATP synthase (ATPase8) shows the most anomalous behavior as both nucleotide and amino acid comparisons give the lowest percentage of similarity. In order to acquire some information on the structure–function relationships and in view of the peculiarity of the sea urchin system, we decided to analyze the ATPase8 gene sequences in these organisms.

ATPase8 is a component of the membrane-integrated F_o sector of the animal mitocchondrial ATP synthase. This enzyme catalyzes the synthesis of ATP during oxidative phosphorylation in a reaction which is driven by an electrochemical gradient generated by the respiratory chain complexes (reviewed by Senior, 1988). The mitochondrial F_o sector, where the proton channel is constituted, has a similar integral membrane subunit composition in different systems. In yeast, detailed analysis of mutants has elucidated the presence of the three mitochondrially encoded subunits 9, 8, and 6 (Law *et al.*, 1990). On the other hand, it is known that in mammalian systems the subunit 9 is encoded in the nuclear genome while the subunit A6L, the animal equivalent of subunit 8, and the subunit 6 are coded by the mitochondrial DNA (Fearnley and Walker, 1986). Although no direct data are available on the composition of the sea urchin F_o sector of mitochondrial ATP synthase, it is conceivable to believe that also in these organisms the F_o sector components are equivalent to those of other animal mitochondrial systems.

In this paper we report detailed comparative studies on the ATP synthase subunit 8 between both closely and distantly related species of echinoderms. The results demonstrate that this gene evolves extremely fast and in a phylum-specific manner.

EXPERIMENTAL PROCEDURE

The entire sequence of the *Arbacia lixula* mtDNA (15719 nt) was obtained in our laboratory (De Giorgi *et al.*, 1996), and submitted to the EMBL data library under accession number X80396. The sequences of *Strongylocentrotus purpuratus* and *Paracentrotus lividus* with accession number X12631 and Y04815 were extracted from database by using the GCG package (Devereux *et al.*, 1984). The other sequences were extracted by using the ACNUC retrieval system (Gouy *et al.*, 1985).

The nucleotide substitution rate was calculated by using the Stationary Markov Model (SMM) method (Saccone *et al.*, 1990). The necessity to use a stochastic model, like the SMM, to do such calculations comes from the possible occurrence of multiple substitutions at the same site as well as reversions. Thus the observed divergence (nucleotide substitutions/site) between two sequences is an underestimate of the actual divergence that can be estimated by using a suitable stochastic model. SMM is one of the most reliable models so far available to calculate the nucleotide divergence between two homologous sequences (Rzhetsky and Nei, 1995).

Multiple alignment was carried out by using the program PILEUP and manually adjusted with LINEUP (GCG package).

The hydropathic profile has been calculated by using the program PEPTIDESTRUCTURE (GCG package) according to the method of Kyte and Doolittle (1982).

The prediction of helical transmembrane regions has been carried out by using the program PHDhtm (Rost *et al.*, 1994). PHDhtm implements a neural network algorithm which is able to learn from a collection of proteins with known transmembrane segments how to single out such segments in other proteins.

RESULTS

By comparison with the other known vertebrate genes so far sequenced, ATPase8 appears to be the least conserved genes in the different mitochondrial systems. However, evidence obtained both in animals (Pedersen and Carafoli, 1987; Walker *et al.*, 1987), but mainly in yeast (Nagley, 1988; Papakonstantinou *et al.*, 1995), have demonstrated that this subunit plays an essential role in the energy transduction by mitochondrial ATP synthase.

Table I. Similarity of ATPase Subunit 8 within the Sea Urchin Genes[a]

	A. lixula	*P. lividus*	*S. purpuratus*
A. lixula	—	61.2	49.7
P. lividus	49.0	—	60.0
S. purpuratus	38.2	44.6	—

[a] The results are indicated in percentage of similarity. Above the dash are reported the values at the nucleotide level; below the dash, those obtained at the amino acid level.

The similarity of the ATPase8 gene within the sea urchins is very low, both at the nucleotide and the amino acid level. The results reported in Table I clearly indicate that the similarity at the amino acid level is even lower than that at the nucleotide level. Furthermore, the observation that the amino acid similarity between *A. lixula* and *P. lividus* is higher than between the two Camarodonta species is also unexpected.

The multialignment of ATPase8 gene in the three species is reported in Fig 1. Identity regions, boxed in the figure, can be detected only at the extremities of the gene. In the central part encompassing roughly the sequence from nucleotide 67 to nucleotide 147, instead, the three sequences diverge, and nucleotide substitutions occur not only between Camarodonta and Stirodonta, but also within the two Camarodonta species.

This unexpected feature prompted us to investigate the dynamics of base substitution with the Stationary Markov Model method (Saccone *et al.*, 1990). It is well known that the substitution rate at the third silent codon positions is higher than at the first and second codon positions because the latter are constrained by the coding requirements. The sea urchin ATPase8 gene, analyzed in a pairwise comparison with the SMM method, has revealed, instead, that the nucleotide substitution rate of nonsynonymous sites (first + second codon positions) is not significantly lower than that of synonymous sites (data not shown), although the large statistical fluctuation due to the small size of the gene does not allow the calculation of precise nucleotide substitution rates. These data clearly indicate a relaxation of functional constraints at the amino acid level. However, when the deduced amino acid sequences were analyzed with a computer-aided method the predicted hydrophatic profiles for the three corresponding proteins are completely superimposable (Fig. 2), thus suggesting that in spite of the divergence observed in the primary structure of the genes in the three species, the role played may be the same.

In Fig. 3A, the multialignment of the deduced amino acid sequences of the sea urchin genes is shown. In order to detect the similarities within the phylum echinoderms, the sequences of two sea star genes (*Asterina pectinifera* and *Pisaster ochraceus*) have also been included. In Fig. 3B, the multialignment of these proteins from different mammalian species is reported. Sea urchin and sea star sequences share a common four-residue sequence motif, (V/M)PQL(X_4)W, at their N-termini, but little if any direct sequence similarity can be detected elsewhere in the proteins. The N-termini of the mammalian sequences (Fig. 3B) also show the same conserved sequence motif MPQL(X_4)W, and in this case several other blocks of conserved amino acids can be found, while the variable

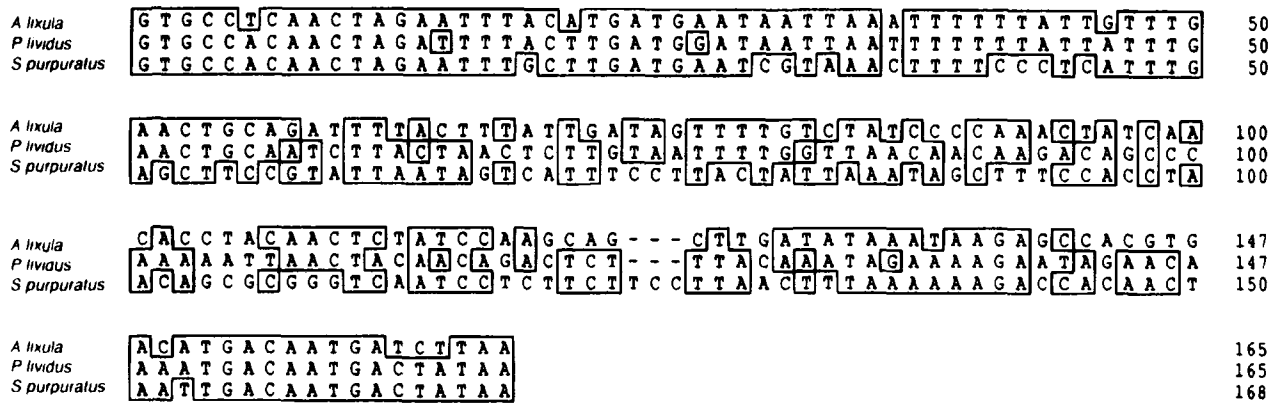

Fig. 1. Multialignment of ATPase8 genes from the sea urchin *Arbacia lixula, Paracentrotus lividus,* and *Strongylocentrotus purpuratus.* Identical nucleotides have been boxed and dots representing gaps have been introduced to optimize the alignment.

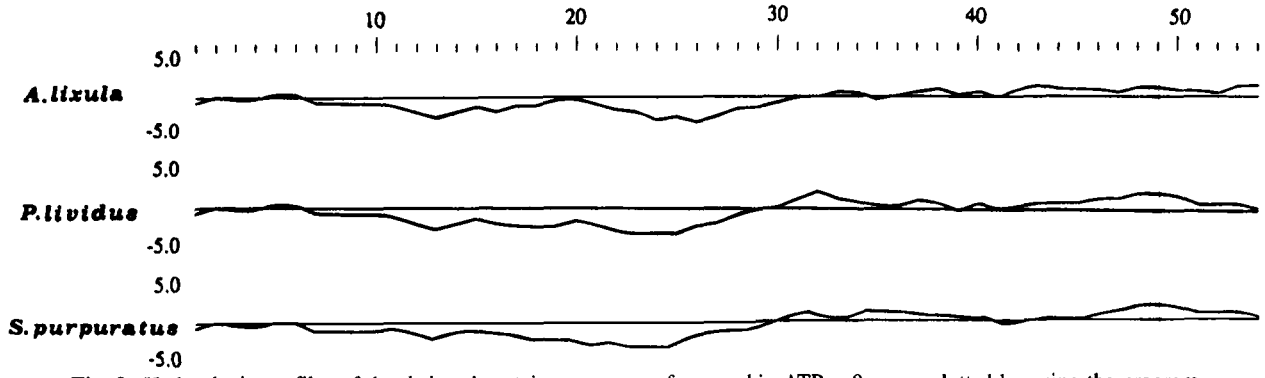

Fig. 2. Hydrophatic profiles of the deduced protein sequences of sea urchin ATPase8 genes plotted by using the program PEPTIDESTRUCTURE (Devereux *et al.*, 1984).

part of the protein encompasses less than 20 out of about 68 amino acid residues.

The results obtained with sea urchin, sea star, and mammals, along with the results obtained in yeast by others (Nagley, 1988), have led us to the conclusion that the sequence motif (V/M)PQL(X_4)(W/F) is evolutionarily conserved from yeast to mammals at the N-termini of the proteins.

Recent advances in the neural network approaches has allowed the developing of computer methods which predict helical trans-membrane regions with more than 95% accuracy (Rost *et al.*, 1994). By using such analysis, the presumed trans-membrane helices have been revealed in the internal left part of the echinoderm sequences (Fig. 3A). This portion is the only segment containing partially conserved amino

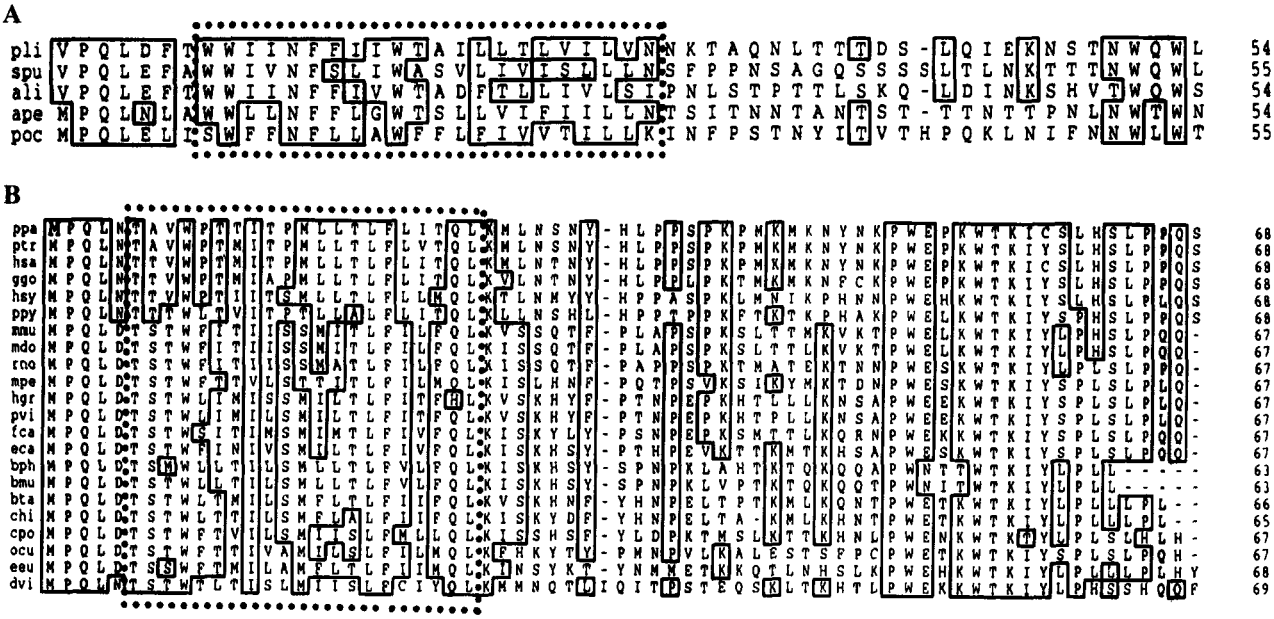

Fig. 3. Multialignment of ATPase8 deduced protein sequences of echinoderms (A) and of mammals (B). The segments included in the area defined by dark dots correspond to the presumed trans-membrane helical regions (Rost *et al.*, 1994). (A) Echinoderm genes: pli: *P. lividus;* spu: *S. purpuratus;* ali: *A. lixula;* ape: *Asterina pectinifera;* poc: *Pisaster ochraceus.* (B) Mammalian genes: ppa: *Pan paniscus;* ptr: *Pan troglodytes;* hsa: *Homo sapiens;* ggo: *Gorilla gorilla;* hsy: *Hylobates syndactylus;* ppy: *Pongo pygmaeus;* mmu: *Mus musculus;* mdo: *Mus domesticus;* rno: *Rattus norvegicus;* mpe: *Microtus pennsylvanicus;* hgr: *Halicoerus grypus;* pvi: *Phoca vitulina;* fca: *Felis catus;* eca: *Equus caballus;* bph: *Balaenoptera physalus;* bmu: *Balaenoptera musculus;* bta: *Bos taurus;* chi: *Capra hircus;* cpo: *Cavia porcellus;* ocu: *Oryctolagus cuniculus;* eeu: *Erinaceus europens;* dvi: *Didelphis virginiana.*

acid residues within the variable domain in the echinoderm system. Similarly, the presumed trans-membrane helices have been localized in the internal left region of the mammalian ATPase8 sequences. The regions which follow the putative trans-membrane domain in mammalian proteins are characterized by the presence of several lysine residues which tend to concentrate toward the C-terminus of the sequence (Fig. 3B). A cluster of positively charged amino acids at the C-terminus of the sequence is a peculiarity described also in yeast (Nagley, 1988). A broad distribution of positively charged amino acid is present also in other mitochondrial systems. In particular, a database search suggests that the bias to accumulate more than one single lysine residue toward the C-terminus is a peculiarity shared by many different organisms. (Table II).

In the case of echinoderms, instead, only the sea urchin species display at the C-terminus a lysine residue at position 47. This lysine residue has been suggested to have some functional role in the yeast protein (Papakonstantinou *et al.*, 1995). At the C-termini of the echinoderm system a peculiar conservation of amino acids has been detected: the motif tryptophan, glutamine, tryptophan, WQW, is conserved in the sea urchin ATPase8 genes, and a very similar motif, W(T/L)W, is conserved also in sea star.

The demonstration that in yeast the F_o sector is assembled *in vivo* through the sequential incorporation of subunits 9,8, and 6 (Hadikusumo *et al.*, 1988) has prompted us to analyze the properties of echinoderm ATP synthase subunit 6 genes. The comparison of the sequence of the subunits 6 within the sea urchin species revealed above 80% identity at the amino acid level while a relevant divergence was observed in the comparison with higher eukaryotes. In particular, the percent identity between echinoderms and humans is on the average 36.8%. On the contrary, a slight difference has been found comparing sea urchins in respect to the corresponding sequences of the sea stars (66.4% identity). Furthermore, as reported in Fig 4, it appears that the echinoderm subunit 6 proteins show the maximum conservation of sequence toward the C-terminus. These observations strongly indicate that also the subunit 6 has evolved in a phylum-specific manner.

Table II. Occurrence of Lysine Residues in ATPase8 Proteins from Different Organisms

Organisms	Number of lysine residues	Database[a] and accession number
Aves:		
Gallus domesticus	3	SW P14093
Amphibia:		
Xenopus laevis	4	SW PO3931
Arthropoda:		
Apis mellifera	7	SW Q00276
Drosophila melanogaster	3	SW P03432
Anopheles gambiae	4	SW P34836
Fishes:		
Cronostoma lacustre	3	SW P34190
Cyprynus carpio	3	SW P24948
Crustacea:		
Artemia franciscana	2	EM X69067
Mollusca:		
Katharina tunicata	3	PIR S50329

[a] SW: Swissport; EM: EMBL; PIR: PIR databases.

DISCUSSION

Recently, great effort has been directed towards an understanding of the mechanism whereby separated components assemble into the multi-subunit enzyme complex of mitchondrial ATP synthase (Papakonstantinou *et al.*, 1995). Evolutionary studies usually play an important role in these investigations. Detailed studies carried out on mutants of *Saccaromyces cerevisiae* indicated that the C-terminal positively charged region of subunit 8 is required for efficient assembly; yeast mutants containing only truncated subunit 8 proteins are defective in the assembly of the F_o sector (Grasso *et al.*, 1991). Therefore, the observation of a broad distribution of positively charged amino acids at the C-terminus of the proteins in many different systems seems to suggest their importance in the assembly of the complex.

The results reported here, instead, clearly indicate the lack of a cluster of positively charged amino acid in the ATPase8 of echinoderms. Furthermore, while few positively charged amino acids are present in these peptides in scattered positions, a lysine residue is conserved in the same position only in sea urchins.

In the mammalian sequences the C-terminus shows a certain degree of variability, with only some amino acids highly conserved. In the case of echinoderms, the conserved amino acids are different in respect to the mammalian counterpart, and a short motif (WQW) seems to be typical. This motif is slightly different also between sea urchin and sea star. Assuming that also in these systems, as in that of yeast, the C-terminus of ATPase8 is involved in the assembly with other components of the F_o sector of ATP synthase, we can conclude that these specific amino acid

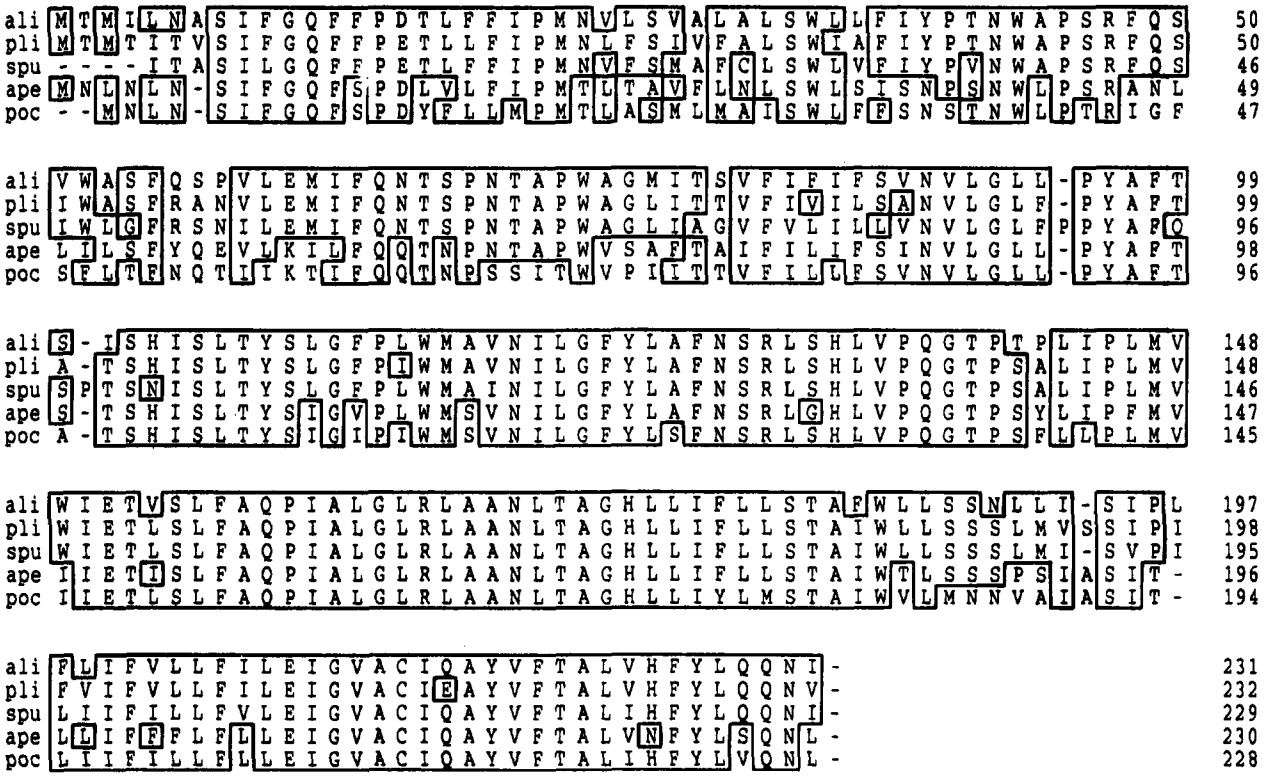

Fig. 4. Multialignment of echinoderm ATPase6 deduced protein sequences. ali: *A. lixula;* pli: *P. lividus;* spu: *S. purpuratus;* ape: *A. pectinifera;* poc: *P. ochraceus.* Identical residues are boxed.

motifs found in echinoderms should play an important role in the interaction with other subunits.

In yeast it has been clearly shown that the mitochondrially coded subunit 9 is the key subunit in the assembly or the stabilization of the F_0 sector. In the case of echinoderms, it seems that the subunit 6 and subunit 8 have co-evolved to facilitate assembly and that signals specific for different taxonomic groups are used in this assembly process.

No data are available regarding the sequences of the nuclear-coded subunit 9 as well as for other nuclear-coded subunits. Therefore, only when this information is available will a complete picture on the mitochondrial-nuclear co-evolution be drawn.

As far as the localization of ATPase8 is concerned, it was commonly accepted that this protein spans the inner mitochondrial membrane once, and is oriented in such way that its N-terminus lies toward the inner-membrane space and its positively charged C-terminus faces the matrix (Velours and Guerin, 1986). This model has been recently disputed and a new hypothesis has been put forward which accommodates the subunit 8 in a hydrophobic niche in contact with other protein subunits of the complex (Papakonstantinou *et al.*, 1995).

Data presented in this paper cannot discriminate between the two possibilities. The observation that the putative transmembrane region has almost an identical size in echinoderms and in mammals, in spite of the different overall length of the proteins, may only indicate that this region is under the same topological constraints.

In conclusion, our results show that during evolution the N-terminus of ATPase8 is conserved in different organisms, where the central part of the molecule undergoes variations but conserves a putative transmembrane segment with a global hydrophobic character, while the C-terminus evolves in a lineage-specific manner.

As for the lineage-specific evolution, it is interesting to recall that, as we have stressed in the Introduction, some features like the gene arrangement remain phylum-specific also in a time span (about 500–600 My) during which the vertebrate–invertebrate split occurred with the generation of a wide variety of mt gene organizations. In addition, it is striking to note

that the amino-acid identity between sea urchin and sea star ATPase6 peptides is much higher (about 67%) than that observed comparing echinoderms and mammals (about 37%) despite, the fact that the time of divergence between sea urchin and sea star (560 My) is comparable with the time of divergence between vertebrates and invertebrates (600 My). This strongly suggest a phylum-specific evolution of the peptides of the ATPase complex which may be misleading for the interphyla comparisons.

These data indicate that at the molecular level there are lineage-specific constraints which should generate detectable "signature" to be used in molecular taxonomy.

ACKNOWLEDGMENTS

This work has been performed with grants from Ministero della Università e della Ricerca Scientifica e Tecnologica, and Human Capital and Mobility Project (EEC grant ERB CH RXCT 930254). The technical assistance of Vito Cataldo is gratefully acknowledged.

REFERENCES

Cantatore, P., Roberti, M., Rainaldi, G., Gadaleta, M. N., and Saccone, C. (1989). *J. Biol. Chem.* **264**, 10965–10975.

Crozier, R. H., and Crozier, Y. C. (1993). *Genetics* **133**, 97–117.

De Giorgi, C., Martiradonna, A., Lanave, C., and Saccone, C. (1995). *Progress in cell Research* (Palmieri, F., Papa, S., Saccone, C., and Gadaleta, M.N., eds.) Elsevier, Amsterdam, pp. 155–159.

De Giorgi, C., Martiradonna, A., Lanave, C., and Saccone, C. (1996). *Mol. Phyl. Evol.* **5**, 323–332.

Devereux, J., Haeberli P., and Smithies, O. (1984). *Nucleic Acids Res.* **12**, 387–395.

Fearnley, I. M., and Walker, J. E. (1986). *EMBO.* **5**, 2003–2008.

Grasso, D. G., Nero, D., Law, R. H. P., Devenish R. J., and Nagley, P. (1991). *Eur. J. Biochem.* **199**, 203–209.

Gouy, M., Gautier, C., Attimonelli, M., Lanave, C., and Di Paola, G. (1985). *CABIOS* **1**, 167–172.

Hadikusumo, R. G., Meltzer, S., Choo, WM., Jean-Francois, M. J. B., Linnane, W., and Marzuki, S. (1988). *Biochim Biophys Acta* **933**, 212–222.

Jacobs, H. T., Elliot, D. J., Math, V. B., Farquharson, A., and Doolittle, R. F. (1982). *J. Mol. Biol.* **157**, 105–132.

Law, R. H. P., Devenish, R. J., and Nagley, P. (1990). *Eur. J. Biochem.* **188**, 421–429.

Nagley, P. (1988). *Trends in Genetics* **4**, 46–52.

Okimoto, R., Macfarlane, J. L., Clary, D. O. and Wolstenholme, D. R. (1992). *Genetics* **130**, 471–498.

Papakonstantinou, T., Law, R. H. P., Manon, S., Devenish, R. J. and Nagley, P. (1995). *Eur. J. Biochem.* **227**, 745–752.

Pedersen, P. L., and Carafoli, E. (1987). *Trends Biochem. Sci.* **12**, 146–150 and 186–189.

Rzhetsky, A., and Nei, M. (1995). *Mol. Biol. Evol.* **12**, 131–151.

Rost, B., Sander, C., and Schneider, R. (1994). *Comput. Appli. Biosci* **10**, 53–60.

Saccone, C., Lanave, C., Pesole, G., and Preparata, G. (1990). *Methods Enzymol.* **183**, 570–583.

Senior, A. E. (1988). *Physiol. Rev.* **68**, 177–231.

Smith, A. B. (1984). *Echinoid Palaeobiology,* Allen and Unwin, London.

Velours, J. and Guerin, B. (1986). *Biochem. Biophys. Res. Commun.* **138**, 78–86.

Walker, J. E., Runswick, M. J., and Poulter, L. (1987). *J. Mol. Biol.* **197**, 89–100.